Coexistence and non-coexistence of Markovian viruses and their hosts

Jakob E. Björnberg, Erik I. Broman

October 30, 2012

Abstract

The possibility of coexistence of two competing populations is a classical question which dates back to the earliest 'predator-prey' models. In this paper we study this question in the context of a model for the spread of a virus infection in a population of healthy cells, introduced in [3]. The infected cells may be seen as a population of 'predators' and the healthy cells as a population of 'prey'. We show that, depending on the parameters defining the model, there may or may not be coexistence of the two populations, and we give precise criteria for this.

1 Introduction

We start by giving an informal description of the model studied in this paper. It is a two-dimensional Markov process $(X(t), Y(t))_{t\geq 0}$, where $X(t)$ is the number of 'healthy cells' at time t, and $Y(t)$ is the number of 'infected cells' (i.e. cells having virus in them). Both components $(X(t))_{t\geq 0}$ and $(Y(t))_{t\geq 0}$ behave in many ways like branching processes, although there are dependencies between them. A healthy cell is replaced by a random number of new healthy cells at rate 1. This random number is independent of other events and drawn from a distribution $(p_k)_{k\geq 0}$; thus the rate at which a healthy cell is replaced by k healthy cells is p_k. Infected cells are also replaced by k new (infected) cells at rate p_k if $k \geq 1$ while they are replaced by 0 new cells (die) at the higher rate $p_0 + \lambda$. Here $\lambda \geq 0$ is a parameter that reflects the negative impact of the virus on the host's lifelength. When an infected cell dies (i.e. is replaced by 0 new cells), it converts a random number of healthy cells into infected cells. The biological motivation is that when infected cells die they burst (lyse) and release 'free virions' which enter a random number of healthy

cells, thus infecting them. The number of conversions is independent of all other events, and is drawn from a distribution $(\gamma_k)_{k\geq 0}$. Hence, the processes $(X(t))_{t\geq 0}$ and $(Y(t))_{t\geq 0}$ interact in that $(Y(t))_{t\geq 0}$ 'feeds' upon $(X(t))_{t\geq 0}$. The model is defined in detail in Section 2. Also, we refer the interested reader to [3] for a biological motivation of the model. We will sometimes simply write X or Y as a shorthand for $(X(t))_{t\geq 0}$ and $(Y(t))_{t\geq 0}$ respectively.

As described, the model is in essence a pair of interacting branching processes. Markov branching processes with interaction have been much studied, see for instance [12] and the references within. The main purpose of this paper is the study of coexistence of the two, competing, populations X and Y. Similar types of questions have been studied in many contexts. One recent example is the so-called two-type-Richardson model. This can be informally described as follows. Consider the graph $\mathbb{Z}^d$, and let the two infections (red and blue) start with only one individual each. A site is infected by the red (blue) process at a rate which equals the infection parameter λ_r (λ_b) times the number of neighbours infected by the red (blue) process. Further, if a site gets infected by the red infection it stays red forever and similarly if it is infected by the blue infection. The main question is if they can coexist, i.e. if there will be two unbounded components of red and blue sites, see for instance [6, 4, 11].

In [3], much focus was on the study of the *extinction probability* η of the infected process $(Y(t))_{t\geq 0}$. There, η was taken as an indicator of the 'evolutionary fitness' of the virus. The main result was that for fixed $(p_k)_{k\geq 0}$ and $(\gamma_k)_{k\geq 0}$ satisfying $\gamma_0 = 0$ the extinction probability η is maximized when $\lambda = 0$. In fact, it was shown that η is increasing in λ. The main result of this paper concerns the *coexistence probability* ζ.

Definition 1.1. We call

$$\zeta = P(X(t)Y(t) \geq 1 \text{ for all } t \geq 0).$$

the coexistence probability of $X(t), Y(t)$.

Of course, ζ depends on the parameters used to define the process, but we suppress this dependence in the notation. Introducing the stopping time $T_{\mathrm{u}} = \inf\{t \geq 0 : X(t)Y(t) = 0\}$, we have that $\zeta = P(T_{\mathrm{u}} = \infty)$. The relevance of coexistence in the study of η will be discussed in Section 6.

The proof of our main result uses two auxiliary branching processes $\hat{X}(t)$ and $\hat{Y}(t)$ defined and discussed in detail in Section 3.1. Informally, $\hat{X}(t)$ is a process distributed as $X(t)$ without the influence of $Y(t)$, i.e. $Y(0) = 0$. Furthermore, $\hat{Y}(t)$ is a process distributed as $Y(t)$ with an infinite supply of healthy cells, i.e. $X(0) = \infty$. Our main result is formulated in terms

of the so-called malthusian parameters for these processes, denoted by α and β for $\hat{X}(t)$ and $\hat{Y}(t)$ respectively (see Section 3.1). It turns out that $\alpha = \sum_{k=0}^{\infty} kp_k - 1$ and $\beta = \alpha + p_0 \sum_{k=0}^{\infty} k\gamma_k + \lambda(\sum_{k=0}^{\infty} k\gamma_k - 1)$.

Theorem 1.2. For arbitrariy initial conditions $X(0), Y(0) \geq 1$ and offspring distributions $(p_k)_{k\geq 0}$ with finite second moment, the coexistence probability is positive if and only if $\alpha > \beta > 0$.

Remark 1.3. Note that coexistence is only possible if $\gamma_0 > 0$, because otherwise $\beta > \alpha$. This follows from the expressions for α and β and Theorem 1.2: if $\gamma_0 = 0$ then Y cannot die out as long as X survives. Hence, there is then almost surely a time $t > 0$ such that $X(t) = 0$ and $Y(t) \neq 0$. (This result was announced in [3] as part of Proposition 3.2.)

Theorem 1.2 establishes, under a second moment condition, for which values of α and β we can have coexistence. Our next result strengthens the second part of Theorem 1.2. Recall that $T_{\mathrm{u}} = \inf\{t > 0 : X(t)Y(t) = 0\}$.

Theorem 1.4. For offspring distributions $(p_k)_{k\geq 0}$ with finite second moment, and for any choice of $\alpha < \beta$ we have that $E[T_{\mathrm{u}}] < \infty$.

We have not been able to establish in general if T_{u} has finite or infinite expectation when $\alpha = \beta$. (But see Remark 5.1 for a special case.)

On the way to proving that coexistence is indeed possible (when $\alpha > \beta$) we use general facts about order statistics and trimmed sums, see Lemma 3.4. The second part of that lemma is an interesting application of Harris' inequality [8] to bound the variance of a trimmed sum, which we have not found in the literature.

In the two-type Richardson model mentioned above, coexistence is conjectured to hold if and only if $\lambda_r = \lambda_b$. The 'if' condition has been established, see [6, 4, 11], while [7] makes progress on the 'only if' condition. In fact, the model studied here is closer to the following variant of the two-type Richardson model. If a site is infected by the blue process, it changes color if the red process attempts to infect it, while if a site is infected by the red process it stays so forever. That is, a red site is immune to the blue process while a blue site is not immune to the red infection. Analogy with the model studied in this paper suggests that there can then be coexistence if $\lambda_b > \lambda_r$, but not if $\lambda_b < \lambda_r$.

We end this section with an outline of the rest of the paper. In Section 2 we give a precise definition of the model. In Section 3 we state and prove preliminary results needed in the proofs of our main results. In Sections 4 and 5 we prove Theorems 1.2 and 1.4 respectively. Finally we discuss some applications of these results in Section 6.

2 Definition

Let $(p_k)_{k\geq 0}$ and $(\gamma_k)_{k\geq 0}$ be probability distributions on the nonnegative integers, and let $\lambda \geq 0$. We exclude the (degenerate) case when $p_1 = 1$; in fact the reader may for convenience assume that $p_1 = 0$, since this only amounts to a time-change.

The continuous–time Markov chain $(X(t), Y(t))_{t\geq 0}$, taking values in $\mathbb{Z}_+^2$, was informally described in Section 1. To recapitulate the main points, each healthy cell is replaced by $k \geq 0$ new healthy cells at rate p_k. Being replaced by $k = 0$ new cells corresponds to dying. Each infected cell is replaced by $k \geq 1$ new infected cells at rate p_k. When an infected cell dies, which occurs at rate $p_0 + \lambda$, a random number of healthy cells are converted into infected cells. If t is the time of such an event, we draw a random variable Γ_t from the distribution $(\gamma_k)_{k\geq 0}$ independently of other events. If $\Gamma_t \leq X(t)$ we simply declare Γ_t of previously healthy cells to be infected, while if $\Gamma_t > X(t)$ we declare all previously healthy cells to be infected. To define this process formally, we list the different possible jumps in Table 1, where we use the notation $x \wedge k = \min\{x, k\}$.

	Transition from (x, y) to	Rate	Valid for
(i)	$(x - 1 + k, y)$	xp_k	$k \geq 0$
(ii)	$(x, y - 1 + k)$	yp_k	$k \geq 1$
(iii)	$(x - (x \wedge k), y - 1 + (x \wedge k))$	$y(p_0 + \lambda)\gamma_k$	$k \geq 0$

Table 1: Transition rates for the process $(X(t), Y(t))_{t\geq 0}$. Rates are given for transitions from a state (x, y) and are valid for all $x, y \geq 0$.

Note that there may be several transitions in Table 1 leading to the same state. In such cases the correct interpretation is to add the corresponding rates. An example of this is the transition $(0, y) \to (0, y - 1)$, which occurs at rate $y(p_0 + \lambda)$, which is the sum over all $k \geq 0$ in (iii). To avoid trivial cases, we assume throughout that $X(0), Y(0) \geq 1$. Biologically it might be most relevant to consider the case when $p_k = 0$ for $k \geq 3$, but all our results are valid in greater generality, so we make no such restriction.

We now state some immediate properties of the model. If it were the case that $Y(t) = 0$, then healthy cells would evolve as a Markov branching process, with intensity 1 and offspring distribution $(p_k)_{k\geq 0}$. Similarly, if $X(t) = 0$ for some t, then $(Y(t+s))_{s\geq 0}$ would behave like a Markov branching process with the higher intensity $(1+\lambda)$ and an offspring distribution derived from $(p_k)_{k\geq 0}$ by placing more mass on $k = 0$. When both $X(t), Y(t) > 0$, as transition rate (iii) tells us, healthy cells may turn into infected cells. This scenario hence 'helps' the process $(Y(t))_{t\geq 0}$ and 'hurts' the process $(X(t))_{t\geq 0}$.

3 Preliminary results

In this section we establish several lemmas which will be used in the proofs of Theorems 1.2 and 1.4. Although their motivation may not be obvious on a first reading, we find it convenient to collect all such preliminary results here so as not to interrupt the flow of the main proofs later.

A note on notation: we will sometimes write a sum of the form $\sum_{k=1}^{a} x_k$ where a is non-integer. The correct interpretation is that the sum goes to the integer part $\lfloor a \rfloor$ but we prefer to omit the $\lfloor \cdot \rfloor$ to keep the notation more readable. A similar comment applies also in other places throughout the paper.

3.1 Auxiliary random variables

It will at several points be useful to compare X and Y to two 'larger' processes $\hat{X}$ and $\hat{Y}$. Here $\hat{X}$ may be thought of as the healthy process in the absence of infection, and $\hat{Y}$ as the infected process in an infinite 'sea' of healthy cells.

To be precise, we let $\hat{X}$ and $\hat{Y}$ be two branching processes with lifelength intensities 1 and $1+\lambda$, and offspring distributions $(p_k)_{k\geq 0}$ and $(q_k)_{k\geq 0}$, respectively, where $(q_k)_{k\geq 0}$ is given by

$$q_0 = \frac{\gamma_0(p_0+\lambda)}{1+\lambda}, \text{ and } q_k = \frac{p_k + \gamma_k(p_0+\lambda)}{1+\lambda} \text{ for } k \geq 1. \tag{1}$$

In Table 2 we give a list of the rates used for the coupling of $(\hat{X}, \hat{Y})$ to (X, Y). However, before that, we give an intuitive explaination.

We start with equal sizes, $\hat{X}(0) = X(0)$ and $\hat{Y}(0) = Y(0)$. Each individual in $X(0)$ is paired with a unique 'friend' in $\hat{X}(0)$, and each individual in $Y(0)$ is paired with a unique friend in $\hat{Y}(0)$. Whenever a cell in X either multiplies or dies a natural death (transition (i) in Table 1) then its friend in $\hat{X}$ undergoes the exact same transition, and the offspring are paired in the natural way. Similarly, whenever a cell in Y multiplies (transition (ii) in Table 1) then its friend in $\hat{Y}$ undergoes the exact same transition, and again the offspring are paired in the natural way. When a cell Y has a lysis (transition (iii) in Table 1), sample a random variable Γ with distribution $(\gamma_k)_{k\geq 0}$. Infect $\Gamma \wedge X$ cells from X, but let the friends in $\hat{X}$ of the newly infected cells in X remain unchanged (but lose their friends, existing as singletons). Proceed by letting the friend in $\hat{Y}$ of the cell in Y which underwent lysis be replaced by Γ new cells. Finally, pair the newly infected cells, now belonging to Y, with the new cells of $\hat{Y}$. Note that if $\Gamma > X$, the this will result in some of the cells in $\hat{Y}$ being unpaired.

Thus every element of X always has a friend in $\hat{X}$, and every element of Y always has a friend in $\hat{Y}$; but some cells in $\hat{X}$ and $\hat{Y}$ might be unpaired. We let unpaired cells give rise to independent Markov branching processes with the correct intensities and offspring distributions. The rates of the coupled process $(X, \hat{X}, Y, \hat{Y})$ are summarized in Table 2. As before, the

Transition from $(x, \hat{x}, y, \hat{y})$ to state	Rate	Valid for
$(x-1+k, \hat{x}-1+k, y, \hat{y})$	xp_k	$k \geq 0$
$(x, \hat{x}-1+k, y, \hat{y})$	$(\hat{x}-x)p_k$	$k \geq 0$
$(x, \hat{x}, y-1+k, \hat{y}-1+k)$	yp_k	$k \geq 1$
$(x, \hat{x}, y, \hat{y}-1+k)$	$(\hat{y}-y)p_k$	$k \geq 1$
$(x-(x \wedge k), \hat{x}, y-1+(x \wedge k), \hat{y}-1+k)$	$y(p_0+\lambda)\gamma_k$	$k \geq 0$
$(x, \hat{x}, y, \hat{y}-1+k)$	$(\hat{y}-y)(p_0+\lambda)\gamma_k$	$k \geq 0$

Table 2: Transition rates in the coupled chain $(X, \hat{X}, Y, \hat{Y})$. Rates are given for transitions from a state $(x, \hat{x}, y, \hat{y})$ and are valid for all $x, \hat{x}, y, \hat{y} \geq 0$. Note that the ordering $x \leq \hat{x}$, $y \leq \hat{y}$ is preserved.

correct interpretation is to add the rates of transitions leading to the same state. We note that our coupling satisfies the following:

1. $X(t) \leq \hat{X}(t)$ and $Y(t) \leq \hat{Y}(t)$ for all $t \geq 0$;
2. if $X(t) \neq 0$ then $\hat{Y}(t) = Y(t)$.

For a probability vector $\pi = (\pi_k : k \geq 0)$ we write $\bar{\pi}$ for the mean $\sum_{k \geq 0} k\pi_k$. Let α and β be the *Malthusian parameters* of $\hat{X}$ and $\hat{Y}$, respectively, given by

$$\alpha = \bar{p} - 1, \qquad \beta = (\bar{q} - 1)(1 + \lambda).$$

It is well known [9] that $\hat{X}(t)/e^{\alpha t}$ and $\hat{Y}(t)/e^{\beta t}$ are martingales which converge almost surely to some nonnegative random variables. We have that $P(A \cup B) = 1$ where

$$A = \big\{\hat{X}(t) = 0 \text{ for some } t \geq 0\big\} \text{ and } B = \big\{\liminf_{t \to \infty} \log(\hat{X}(t))/t > 0\big\}. \quad (2)$$

Moreover, $P(A) = 1$ if and only if $\alpha \leq 0$. On the event B, the limit $\lim_{t \to \infty} \log(\hat{X}(t))/t$ exists and equals α. The corresponding statements hold for $\hat{Y}(t)$ with α replaced by β. Note for future reference that

$$\begin{aligned} \beta = (\bar{q} - 1)(1 + \lambda) &= \Big(\frac{\bar{p} + \bar{\gamma}(p_0 + \lambda)}{1 + \lambda} - 1\Big)(1 + \lambda) \\ &= \bar{p} - 1 + p_0\bar{\gamma} + \lambda(\bar{\gamma} - 1) = \alpha + p_0\bar{\gamma} + \lambda(\bar{\gamma} - 1). \end{aligned} \quad (3)$$

Next, let U, V, W and Φ denote random variables with the following distributions. Firstly, U and V have the distributions of (the sizes of) $\hat{X}(1)$ and $\hat{Y}(1)$, respectively, when $\hat{X}(0) = 1$ and $\hat{Y}(0) = 1$. Secondly, W has the distribution of $\tilde{X}(1)$, where $\tilde{X}$ is a branching process, started at 1, with lifelength intensity 1 and offspring distribution π given by $\pi_0 = 0$, $\pi_1 = p_0 + p_1$, and $\pi_k = p_k$ for $k \geq 2$. Thus $\tilde{X}$ is essentially $\hat{X}$ with deaths suppressed. Finally, to define Φ run a sample of $\hat{Y}$ for time 1, started with $\hat{Y}(0) = 1$; for each branching event that occurs during this time sample an independent Bernoulli random variable with success probability p_0, and let L denote the total number of successes. Let Φ have the distribution of a sum of L independent copies of Γ. Thus Φ is, intuitively, the number of infection attempts during a unit time interval starting with one infected cell.

Lemma 3.1. Let $r \geq 1$ and let D denote a random variable with distribution $(p_k)_{k\geq 0}$. Then

1. $E(U^r) < \infty$ if $E(D^r) < \infty$,
2. $E(V^r) < \infty$ if $E(D^r) < \infty$ and $E(\Gamma^r) < \infty$,
3. $E(W^r) < \infty$ if $E(D^r) < \infty$,
4. $E(\Phi) < \infty$ if $E(\Gamma) < \infty$ and $E(D) < \infty$.

Proof. From [2, Corollary III.6.1], we know that a branching process with offspring distribution π has finite rth moment at time $t > 0$ if π has its rth moment. This immediately gives parts 1 and 3. Part 2 follows from (1), which implies that $(q_k)_{k\geq 0}$ has its rth moment if $(p_k)_{k\geq 0}$ and $(\gamma_k)_{k\geq 0}$ do. For the final part, note that

$$E(\Phi) = E\Big(\sum_{j\geq 1} \Gamma_j \mathbb{1}\{L \geq j\}\Big) = E(\Gamma)E(L).$$

An easy (stochastic) upper bound on L is given by $\tilde{Y}(1)$ where $\tilde{Y}$ is a branching process with intensity $1 + \lambda$ and offspring distribution $(\tilde{q}_k)_{k\geq 0}$, where $\tilde{q}_0 = \tilde{q}_1 = 0$, $\tilde{q}_2 = q_2 + q_1 + q_0$, and $\tilde{q}_k = q_k$ for $k \geq 3$. Thus $E(L)$ is finite if $E(\Gamma)$ and $E(D)$ are finite, as in part 2. □

We will in what follows always assume that $(p_k)_{k\geq 0}$ has finite second moment, since this is part of the assumptions in Theorems 1.2 and 1.4. By Lemma 3.1 this implies that $E(U^2) < \infty$, $E(V^2) < \infty$, $E(W^2) < \infty$ and $E(\Phi) < \infty$. This will allow us to apply Chebyshev's bound, which we will use in the following form. Let Z_j $(j \geq 1)$ be independent, all with the same

nonnegative mean $E(Z) \geq 0$ and finite variance $\mathrm{Var}(Z) < \infty$ as some random variable Z. Let $N \geq 1$ be any integer and let $\delta > 0$. Then

$$\begin{aligned} P\Big(\sum_{j=1}^{N} Z_j > (1+\delta)NE(Z)\Big) &\leq P\Big(\Big[\sum_{j=1}^{N} Z_j - E(Z_j)\Big]^2 > N^2\delta^2 E(Z)^2\Big) \\ &\leq \frac{N\mathrm{Var}(Z)}{N^2\delta^2 E(Z)^2} = \frac{1}{N}\cdot\frac{\mathrm{Var}(Z)}{\delta^2 E(Z)^2}. \end{aligned} \tag{4}$$

Similarly

$$P\Big(\sum_{j=1}^{N} Z_j < (1-\delta)NE(Z)\Big) \leq \frac{1}{N}\cdot\frac{\mathrm{Var}(Z)}{\delta^2 E(Z)^2}. \tag{5}$$

3.2 Estimates

The following lemma says that Y cannot be much larger than X for very long without making X extinct. This lemma will be the main step in the proof of the case $\beta > \alpha$ in Theorem 1.2, which is the case when the process $\hat{Y}$ grows much faster than X. In the statement of the lemma, we let W be as in Lemma 3.1, and let ξ be a Bernoulli variable with success probability $1 - e^{-(1-\gamma_0)(p_0+\lambda)}$ (this being the probability of a lysis leading to at least one new infection occuring in a time interval of length 1). We fix $c > 0$ and let $\delta(t) > 0$ be any function such that

$$n\delta(n) > \frac{1}{2}\log\left(2\frac{E(W)}{E(\xi)}\right) \tag{6}$$

for all sufficiently large n. We write

$$A_n = \{\forall t \in [n, n+1], 0 < X(t) \leq e^{(c-\delta(t))t} < e^{(c+\delta(t))t} \leq Y(t)\}.$$

Lemma 3.2. There is a constant $C > 0$ such that for n large enough that (6) holds,

$$P(A_n) \leq Ce^{-(c-\delta(n))n}. \tag{7}$$

In particular, we can take $C = 9(\mathrm{Var}(W)/E(W)^2 + \mathrm{Var}(\xi)/E(\xi)^2)$. It follows that $P(A_n \text{ i.o.}) = 0$.

Before turning to the proof we remark that we only actually use this lemma with δ constant. We prove this slightly more general result since very little extra work is required, and we hope that it will be useful for future work.

Proof. The result is trivial if $\delta(n) \geq c$ so we assume that $\delta(n) < c$; we also assume throughout the proof that n is large enough that (6) holds. Suppose that A_n occurs. Let Φ_n denote the number of infection attempts during the time interval $[n, n+1]$, that is to say the sum of an independent sample of Γ for each lysis of $(Y(t) : t \in [n, n+1])$. Let $\xi^{(n)}$ be obtained from $(Y(t) : t \in [n, n+1])$ as follows. Start by numbering the elements of $Y(n)$ (arbitrarily); then observe those elements numbered at most $e^{(c+\delta(n))n}$ until they undergo a branching event; let ξ_j be the indicator of the event that cell j has a branching event which results in a lysis for which the associated Γ-value is at least 1 ($\xi_j = 0$ if there is no branching event before time $n+1$); finally let $\xi^{(n)}$ be the sum of the ξ_j. Then $\xi^{(n)}$ has the following properties:

1. $\xi^{(n)} \leq \Phi_n$,
2. $\xi^{(n)}$ is a sum of $e^{(c+\delta(n))n}$ independent Bernoulli variables, each with success probability $1 - e^{-p_0(1-\gamma_0)(1+\lambda)}$, and
3. $\xi^{(n)}$ is independent of $(X(t) : t \in [n, n+1])$.

Next, let $W^{(n)}$ denote the total number of healthy cells that ever exist in the time-interval $[n, n+1]$. Of course, if $W^{(n)} \leq \Phi_n$, then A_n cannot occur since this would imply that $X(n+1) = 0$. We cannot *immediately* conclude from the fact that $X(t) \leq e^{(c-\delta(t))t}$ for every $t \in [n, n+1]$, that $W^{(n)}$ is bounded by $e^{(c-\delta(n+1))(n+1)}$. However $W^{(n)}$ must be stochastically bounded by the sum of $e^{(c-\delta(n))n}$ independent copies W_j of the random variable W in Lemma 3.1. (Recall that W is, intuitively, $\hat{X}(1)$ when deaths are suppressed.) Also, $W^{(n)}$ is independent of $\xi^{(n)}$. Thus, writing $a_n = e^{(c-\delta(n))n}$ and $b_n = e^{(c+\delta(n))n}$, we have that

$$P(A_n) \leq P(W^{(n)} > \xi^{(n)}) \leq P\Big(\sum_{j=1}^{a_n} W_j > \sum_{j=1}^{b_n} \xi_j\Big)$$
$$= P\Big(\frac{1}{a_n}\sum_{j=1}^{a_n} W_j > \frac{b_n}{a_n}\frac{1}{b_n}\sum_{j=1}^{b_n} \xi_j\Big).$$

Note that $b_n/a_n = e^{2n\delta(n)} > e^{\log(2E(W)/E(\xi))} = 2E(W)/E(\xi)$, by (6). We get

that

$$\begin{aligned} P(A_n) &\le P\Big(\frac{1}{a_n}\sum_{j=1}^{a_n} W_j > 2\frac{E(W)}{E(\xi)}\frac{1}{b_n}\sum_{j=1}^{b_n}\xi_j\Big) \\ &\le P\Big(\frac{1}{a_n}\sum_{j=1}^{a_n} W_j > 2\frac{E(W)}{E(\xi)}\frac{2}{3}E(\xi)\Big) + P\Big(\frac{2}{3}E(\xi) > \frac{1}{b_n}\sum_{j=1}^{b_n}\xi_j\Big) \\ &\le \frac{9\mathrm{Var}(W)}{a_n E(W)^2} + \frac{9\mathrm{Var}(\xi)}{b_n E(\xi)^2}, \end{aligned}$$

where we use (4) and (5). This gives (7). That $P(A_n \text{ i.o.}) = 0$ follows from the Borel–Cantelli lemma. □

Recall that if $U(t)$ is a Markov branching process with Malthusian parameter u then $W(t) = U(t)/e^{ut}$ is a martingale. We make no claim as to the originality of the following lemma, yet have not seen it explicitly formulated.

Lemma 3.3. Let $U(t)$ be a branching process whose offspring distribution has finite second moment and with Malthusian parameter $u > 0$.

1. For any $\Delta > 0$ we have that

$$P(\exists t \ge 0 : W(t) \ge \Delta) \le \Delta^{-1}.$$

2. For each $\varepsilon > 0$ there is some $\kappa > 0$ such that

$$P(\exists t \ge \tau : 0 < W(t) < e^{-\varepsilon t}) \le e^{-\kappa\tau}.$$

Proof. The first part is simply a consequence of Doob's submartingale inequality, which gives that for any $T > 0$,

$$P(\exists t \in [0,T] : W(t) \ge \Delta) = P\Big(\sup_{0\le t\le T} W(t) \ge \Delta\Big) \le E[W(T)]/\Delta = \Delta^{-1}.$$

Letting $T \to \infty$ concludes the proof of this case.

For the second part, we proceed by discretizing. Let $\mu = E[U(1)] = e^u$ and let $W_n = U(n)/\mu^n$ for every $n \in \mathbb{N}$. It is no loss of generality to assume that $\varepsilon < u/2$. The limit $W := \lim_n W_n$ exists a.s. since $(W_n)_{n\ge1}$ is a nonnegative martingale. A straightforward and standard calculation (see for instance [9, p. 13]) shows that for any $r > n$,

$$E[\mu^n(W_r - W_n)^2] = \sigma^2(\mu^{-1} + \mu^{-2} + \cdots + \mu^{-r}),$$

where $\sigma^2 = \mathrm{Var}(U(1))$. Therefore by Fatou's lemma

$$E[(W - W_n)^2] \le \liminf_{r\to\infty} E[(W_r - W_n)^2] = \frac{\sigma^2}{\mu - 1}\mu^{-n} \tag{8}$$

for all n. Hence by Markov's inequality

$$P(|W - W_n| > e^{-\varepsilon n}) \le \frac{E[(W - W_n)^2]}{e^{-2\varepsilon n}} \le \frac{\sigma^2}{\mu - 1} e^{-(u-2\varepsilon)n}.$$

It is well known (see for instance [9, Theorem 8.3]) that there exists a constant $c_3 > 0$ such that for any interval $I \subset (0, \infty)$ we have $P(W \in I) \le c_3|I|$. Furthermore, it is also well known [9, Theorem 8.4] that there exists a constant $c_4 > 0$ such that $P(W = 0, W_n \ne 0) \le e^{-c_4 n}$. Therefore (adjusting c_3 as necessary)

$$\begin{aligned} P(0 < W_n < e^{-\varepsilon n}) &\le P(W = 0, W_n > 0) + P(0 < W < 2e^{-\varepsilon n}) \\ &\quad + P(|W - W_n| > e^{-\varepsilon n}) \\ &\le c_3(e^{-c_4 n} + e^{-\varepsilon n} + e^{-(u-2\varepsilon)n}). \end{aligned} \tag{9}$$

Clearly

$$\begin{aligned} P(\exists s \ge t : 0 < W(s) < e^{-\varepsilon s}) &\le P(\exists n \ge t : 0 < W_n < e^{-\varepsilon n/2}) \\ &\quad + P(\exists s \ge t : 0 < W(s) < e^{-\varepsilon s}, \forall n \ge t \; W_n = 0 \text{ or } W_n \ge e^{-\varepsilon n/2}). \end{aligned} \tag{10}$$

We have bounded the first probability on the right hand side in (9). The second probability is bounded above by

$$\begin{aligned} &P\Big(\bigcup_{n\ge t}\{\exists s \in [n, n+1] : W(s) < e^{-\varepsilon n},\ W_n \ge e^{-\varepsilon n/2}\}\Big) \\ &\quad \le \sum_{n\ge t} P(\exists s \in [n, n+1] : W(s) < e^{-\varepsilon n} \mid W_n \ge e^{-\varepsilon n/2}) P(W_n \ge e^{-\varepsilon n/2}) \\ &\quad \le \sum_{n\ge t} P(\exists s \in [n, n+1] : U(s) < e^{u(n+1)-\varepsilon n} \mid U(n) \ge e^{un - \varepsilon n/2}). \end{aligned}$$

It therefore suffices to show that each of the summands is exponentially small in n for large enough n.

To establish this we take the following point of view. Let $M = U(n)$ and label the particles present at time n by $1, 2, \ldots, M$. If particle j has a branching event with zero offspring we say that particle j is *destroyed*. If it has a branching event with one or more offspring, we consider particle j to be still present, essentially identifying it with one of its offspring particles. With

this convention, we let A_j denote the event that particle j is ever destroyed during the time interval $[n, n+1]$. Thus $P(A_j) < 1$ for all j, and the events A_j are independent. If $U(s) \le e^{u(n+1)-\varepsilon n}$ for some $s \in [n, n+1]$ then at least $M - e^{u(n+1)-\varepsilon n}$ of the events A_j must occur. But since $M \ge e^{un-\varepsilon n/2}$

$$P\Big(\sum_{j=1}^{M} \mathbb{1}_{A_j} \ge M - e^{u(n+1)-\varepsilon n}\Big) \le P\Big(\sum_{j=1}^{M} \mathbb{1}_{A_j} \ge M(1 - e^u e^{-\varepsilon n/2})\Big)$$
$$\le P\Big(\sum_{j=1}^{M} \mathbb{1}_{A_j} \ge MP(A_j)(1+\delta)\Big)$$

for large enough n and some $\delta > 0$. The latter probability is by (4) at most

$$C/M \le Ce^{-(u-\varepsilon/2)n}.$$

This gives the result. □

3.3 A lemma about order statistics

The following result will be used in the case $\alpha > \beta$ in Theorem 1.2, but may also be of independent interest. The first part essentially goes back to [1] (in the case $p = 2$), but we have not found the second part in the literature.

If $(X_j)_{1\le j\le M}$ is a sequence of indentically distributed random variables, we let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(M)}$ denote the *order statistics* of $(X_j)_{1\le j\le M}$.

Lemma 3.4. Let $(X_j)_{1\le j\le M}$ be as above.

1. If $p > 1$ and $\|X_1\|_p = E[X_1^p]^{1/p} < \infty$ then for each subset $A \subseteq \{1, \dots, M\}$,
$$E\Big[\sum_{j\in A} X_{(j)}\Big] \le \|X_1\|_p M^{1/p} m^{1/q}, \tag{11}$$
where $m = |A|$ and $1/p + 1/q = 1$.

2. If the X_i are independent and $E[X_1^2] < \infty$, then
$$\mathrm{Var}\Big(\sum_{j=1}^{M-m} X_{(j)}\Big) \le \mathrm{Var}\Big(\sum_{j=1}^{M} X_j\Big) = M \cdot \mathrm{Var}(X_1). \tag{12}$$

Proof. The first part is a consequence of Hölder's inequality:

$$E\Big[\sum_{j\in A} X_{(j)}\Big] = E\Big[\sum_{j=1}^{M} X_{(j)} \mathbb{1}\{j \in A\}\Big] \le E\Big[\sum_{j=1}^{M} |X_{(j)}|^p\Big]^{1/p} E\Big[\sum_{j=1}^{M} \mathbb{1}\{j \in A\}\Big]^{1/q}$$
$$= E\Big[\sum_{j=1}^{M} |X_j|^p\Big]^{1/p} |A|^{1/q} = \|X_1\|_p m^{1/q} M^{1/p}.$$

For the second part, let X denote the sequence $(X_1, \dots, X_M)$ and let

$$f(X) = \sum_{j=1}^{M-m} X_{(j)} \quad \text{and} \quad g(X) = \sum_{j=M-m+1}^{M} X_{(j)}.$$

Note that both f and g are increasing functions in the sense that if $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ satisfy $x_i \le y_i$ for every $i = 1, \dots, n$, then $f(x) \le f(y)$ and $g(x) \le g(y)$. Thus also $f(X) - E[f(X)]$ and $g(X) - E[g(X)]$ are increasing functions. It follows from Harris' inequality that

$$\begin{aligned} &E\big[(f(X) - E[f(X)])(g(X) - E[g(X)])\big] \\ &\qquad \ge E\big[f(X) - E[f(X)]\big] E\big[g(X) - E[g(X)]\big] = 0, \end{aligned}$$

that is to say

$$\mathrm{Cov}\Big(\sum_{j=1}^{M-m} X_{(j)}, \sum_{j=M-m+1}^{M} X_{(j)} \Big) \ge 0.$$

It follows that

$$\begin{aligned} \mathrm{Var}\Big(\sum_{j=1}^{M} X_j \Big) &= \mathrm{Var}\Big(\sum_{j=1}^{M-m} X_{(j)} + \sum_{j=M-m+1}^{M} X_{(j)} \Big) \\ &= \mathrm{Var}\Big(\sum_{j=1}^{M-m} X_{(j)} \Big) + \mathrm{Var}\Big(\sum_{j=M-m+1}^{M} X_{(j)} \Big) \\ &\quad + 2\mathrm{Cov}\Big(\sum_{j=1}^{M-m} X_{(j)}, \sum_{j=M-m+1}^{M} X_{(j)} \Big) \\ &\ge \mathrm{Var}\Big(\sum_{j=1}^{M-m} X_{(j)} \Big). \end{aligned}$$

□

Setting $m = 1$ in (11) we deduce that $E[X_{(M)}]$ is of order at most $M^{1/p}$ when the X_i have finite p:th moment. Results of this type, usually formulated for $p = 2$, go back to [1, 5, 10]

Note that (11) is in some sense sharpest when $A = \{M - m + 1, \dots, M\}$ because then the sum consists of the m largest terms; this is the case we will be using.

4 Proof of Theorem 1.2

Clearly (by (2)) $\zeta = 0$ if either $\alpha \leq 0$ or $\beta \leq 0$, so we assume henceforth that $\alpha, \beta > 0$. The proof of Theorem 1.2 will be divided into the three cases (i) $\alpha < \beta$, (ii) $\alpha = \beta$ and (iii) $\alpha > \beta$.

The case $\alpha < \beta$. The intuition is that if coexistence were to take place, then $Y(t)$ would eventually be *much* larger than $X(t)$; but then there is a good chance that all healthy cells are infected in, say, time 1, which would contradict coexistence. To make this intuition exact, let $c = (\alpha + \beta)/2$, $\delta = (\beta - \alpha)/4 > 0$, and use (2) to see that

$$\begin{aligned} P(T_{\mathrm{u}} = \infty) &= P(T_{\mathrm{u}} = \infty, \exists t_0 : \hat{X}(t) \leq e^{(c-\delta)t} < e^{(c+\delta)t} \leq \hat{Y}(t),\ \forall t \geq t_0) \\ &\leq P(\exists t_0 : 0 < X(t) \leq e^{(c-\delta)t} < e^{(c+\delta)t} \leq Y(t),\ \forall t \geq t_0), \end{aligned} \tag{13}$$

since on $\{T_{\mathrm{u}} = \infty\}$ we have that $0 < X(t) \leq \hat{X}(t)$ and $Y(t) = \hat{Y}(t)$ for every $t \geq 0$. Trivially, the right hand side is bounded above by $P(A_n \text{ i.o.})$ where

$$A_n := \{\forall t \in [n, n+1], 0 < X(t) \leq e^{(c-\delta)t} < e^{(c+\delta)t} \leq Y(t)\}.$$

But $P(A_n \text{ i.o.}) = 0$ by Lemma 3.2. □

The case $\alpha = \beta$. For the case $\alpha = \beta > 0$ the intution is that there will typically be so many infection events that X *effectively* (i.e. counting losses due to infections) has a strictly larger rate of deaths than $\hat{X}$, allowing us to essentially reduce this case to the case $\alpha < \beta$. Note that the process

$$R(t) = \frac{\hat{Y}(t)}{\hat{X}(t)} = \frac{\hat{Y}(t)}{e^{\alpha t}} \frac{e^{\alpha t}}{\hat{X}(t)}$$

converges almost surely to some random variable R, since $\hat{Y}(t)/e^{\alpha t}$ and $\hat{X}(t)/e^{\alpha t}$ are nonnegative martingales. The limit R may be infinite, but on the event $\{T_{\mathrm{u}} = \infty\}$ we have that $0 < R < \infty$. Furthermore, since 0 is an absorbing state for the process $R(t)$ we have (up to a null event) that $\{T_{\mathrm{u}} = \infty\} \subseteq \{\inf_{t \geq 0} R(t) > 0\}$. It follows that for each $r > 0$ we have

$$\{T_{\mathrm{u}} = \infty\} \subseteq \{0 < \inf_{t \geq 0} R(t) < r\} \cup G_r,$$

where

$$G_r = \Big\{ \frac{\hat{Y}(t)}{\hat{X}(t)} \geq r \ \forall t \geq 0 \Big\} \cap \{X(t)Y(t) > 0 \ \forall t \geq 0\}.$$

For each $\delta > 0$ we may choose $r > 0$ sufficiently small so that $P(0 < \inf_{t\geq 0} R(t) < r) \leq \delta$ and thus $\zeta \leq \delta + P(G_r)$. We aim to show that $P(G_r) = 0$ for each $r > 0$; since $\delta > 0$ was arbitrary this will complete the proof.

Fix $\delta, r > 0$ as above. We will couple X, $\hat{X}$ and Y to a new process X' which is obtained by taking into account some of the effect of Y on X. The process $X'(t)$ will be a Markov branching process and will satisfy $X'(t) \leq \hat{X}(t)$ for all $t \geq 0$. Moreover, on the event G_r we will have that $X(t) \leq X'(t)$ for all $t \geq 0$. We let $X'(0) = X(0)$. The rates governing the quadruple $(X, X', \hat{X}, Y)$ are given in Table 3, where we have written

$$\kappa = \kappa(x', y) = \left(r\frac{x'}{y}\right) \wedge 1.$$

Transition to state	Rate	For
$(x+k-1, x'+k-1, \hat{x}+k-1, y)$	$(x \wedge x')p_k$	$k \geq 0$
$(x+k-1, x', \hat{x}+k-1, y)$	$(x - x \wedge x')p_k$	$k \geq 0$
$(x, x'+k-1, \hat{x}+k-1, y)$	$(x' - x \wedge x')p_k$	$k \geq 0$
$(x, x', \hat{x}+k-1, y)$	$(\hat{x} - x \vee x')p_k$	$k \geq 0$
$(x, x', \hat{x}, y+k-1)$	yp_k	$k \geq 1$
$(x, x', \hat{x}, y-1)$	$y(p_0+\lambda)\gamma_0$	
$(x-(x\wedge k), x', \hat{x}, y-1+(x\wedge k))$	$y(p_0+\lambda)\gamma_k(1-\kappa)$	$k \geq 1$
$(x-(x\wedge k), x'-1, \hat{x}, y-1+(x\wedge k))$	$y(p_0+\lambda)\gamma_k\kappa$	$k \geq 1$
$(x, x'-1, \hat{x}, y)$	$(rx' - \kappa y)(p_0+\lambda)(1-\gamma_0)$	

Table 3: Transition rates in the coupled chain $(X, X', \hat{X}, Y)$. Rates are given for transitions from a state $(x, x', \hat{x}, y)$ and are valid for all $x, x', \hat{x}, y \geq 0$.

We note from Table 3 that the triple $(X, \hat{X}, Y)$ has the correct marginal distribution, i.e. as described in Section 3.1. For example, summing the first two lines the of the table gives the rate $(x \wedge x' + x - x \wedge x')p_k = xp_k$ for the transition $x \to x+k-1$. Similarly, $\hat{x} \to \hat{x}+k-1$ at rate given by the sum of the first four lines, and using that $x + x' - x \wedge x' = x \vee x'$ we get the correct rate $\hat{x}p_k$.

Consider now the marginal distribution for X'. First note that, since $\kappa \leq rx'/y$, the final rate in Table 3 is nonnegative. Adding the rates for the transitions $x' \to x' - 1$, we find that this transition occurs at rate

$$x'(p_0 + r(p_0+\lambda)(1-\gamma_0)).$$

Together with the rates for $x' \to x' + k - 1$ for $k \geq 1$, this means that $X'(t)$ is a Markov branching process with lifelength intensity

$$1 + r(p_0+\lambda)(1-\gamma_0)$$

and offspring distribution p' given by

$$p'_0 = \frac{p_0 + r(p_0+\lambda)(1-\gamma_0)}{1 + r(p_0+\lambda)(1-\gamma_0)},$$
$$p'_k = \frac{p_k}{1 + r(p_0+\lambda)(1-\gamma_0)}, \quad k \geq 1.$$

In particular, the Malthusian parameter of X' is

$$\begin{aligned}\alpha' &= (1 + r(p_0+\lambda)(1-\gamma_0))(\bar{p'} - 1) \\ &= \alpha - r(p_0+\lambda)(1-\gamma_0) \\ &< \alpha \quad \text{for } r > 0.\end{aligned}$$

Clearly $X'(t) \leq \hat{X}(t)$ for all $t \geq 0$. On the event G_r we also have that $\hat{Y}(t) = Y(t)$ for all $t \geq 0$ and that $\hat{Y}(t)/\hat{X}(t) \geq r$ for all $t \geq 0$. It follows that, on G_r, we have that

$$r \leq \frac{\hat{Y}(t)}{\hat{X}(t)} = \frac{Y(t)}{\hat{X}(t)} \leq \frac{Y(t)}{X'(t)} \quad \text{for all } t \geq 0,$$

so that $rX'(t)/Y(t) \leq 1$ and hence $\kappa(X'(t), Y(t)) = rX'(t)/Y(t)$. Thus the final rate in Table 3 is always 0 on the event G_r, and hence so is the second rate. Therefore, we get that $G_r \subseteq \{X(t) \leq X'(t)\, \forall t \geq 0\}$.

Let $c = (\alpha+\alpha')/2$ and $\delta = (\alpha-\alpha')/4 > 0$. Using (2) we therefore deduce that

$$P(G_r) \leq P(\exists t_0 : 0 < X(t) \leq X'(t) \leq e^{(c-\delta)t} < e^{(c+\delta)t} \leq Y(t),\ \forall t \geq t_0). \tag{14}$$

By Lemma 3.2, the probability on the right equals zero. Since $\delta > 0$ was arbitrary it follows that $\zeta = 0$. □

We are now ready to prove the final case of Theorem 1.2.

The case $\alpha > \beta$. The intuition here is that $X(t)$ 'wants' to be of the order $e^{\alpha t}$ and $Y(t)$ 'wants' to be of the, much smaller, order $e^{\beta t}$. Typically, therefore, the infection will have very little impact on the healthy population.

To make this intuition rigorous, let

$$a_n = \prod_{k=2}^{n} \Big(1 - \frac{2}{k^2}\Big), \quad b_n = \prod_{k=2}^{n} \Big(1 - \frac{1}{k^2}\Big), \quad c_n = \prod_{k=2}^{n} \Big(1 + \frac{1}{k^2}\Big).$$

Note that a_n and b_n form decreasing sequences with limits in $(0,1)$ and that c_n is an increasing sequence with limit in $(1,\infty)$. Write B_n for the event that

$$X(n) \geq a_n e^{\alpha n} \text{ and } b_n e^{\beta n} \leq Y(n) \leq c_n e^{\beta n}.$$

We will prove that there is some N such that

$$P(B_{n+1} \mid B_n) \geq 1 - \frac{3}{n^2} \tag{15}$$

for all $n \geq N$. This will, using the Markov property, establish the result, since $P(B_N) > 0$ and

$$\zeta \geq P(\cap_{n \geq N} B_n) = P(B_N) \prod_{n \geq N} P(B_{n+1} \mid B_n) > 0.$$

We start by observing that (again using that $Y(t) = \hat{Y}(t)$ whenever $X(t) > 0$)

$$\begin{aligned} &P(B_{n+1} \mid B_n) \\ &\quad = P(X(n+1) \geq a_{n+1} e^{\alpha(n+1)}, b_{n+1} e^{\beta(n+1)} \leq \hat{Y}(n+1) \leq c_{n+1} e^{\beta(n+1)} \mid B_n) \\ &\quad \geq 1 - P(X(n+1) < a_{n+1} e^{\alpha(n+1)} \mid B_n) - P\big(\hat{Y}(n+1) < b_{n+1} e^{\beta(n+1)} \mid B_n\big) \\ &\qquad - P\big(\hat{Y}(n+1) > c_{n+1} e^{\beta(n+1)} \mid B_n\big). \end{aligned} \tag{16}$$

We will proceed to show that all three probabilities on the right hand side are small. To prove that $P(X(n+1) < a_{n+1} e^{\alpha(n+1)} \mid B_n)$ is small, let Φ_n denote the number of infection attempts during the time interval $[n, n+1]$, as in the proof of Lemma 3.2. We will first show that Φ_n will typically be much smaller than $X(n)$, and will deduce from this the required lower bound on $X(n+1)$. For the bound on Φ_n, we use Markov's inequality to see that

$$P(\Phi_n \geq c_n e^{\beta n} \cdot (n+1)^2 E(\Phi) \mid B_n) \leq \frac{E(\Phi_n \mid B_n)}{c_n e^{\beta n} \cdot (n+1)^2 E(\Phi)} \leq \frac{1}{(n+1)^2}, \tag{17}$$

where Φ is the random variable of Lemma 3.1 and we used the fact that, given B_n, the number Φ_n of infection attempts is dominated by the sum of $c_n e^{\beta n}$ independent copies of Φ.

Let $M = M(n) = a_n e^{\alpha n}$ and $m = m(n) = c_n e^{\beta n}(n+1)^2 E(\Phi)$ (so $X(n) \geq M$ on B_n, and m is the quantity in (17)). Let $(U_j)_{1 \leq j \leq M}$ denote independent copies of the random variable U of Lemma 3.1. The lower bound on $X(n+1)$ will be obtained by noting that the impact of infection during the time interval $[n, n+1]$ can be no larger than the effect of removing, at time n, those Φ_n healthy cells that would otherwise give rise to the largest ancestry at time $n+1$. In particular,

$$X(n+1) \geq \sum_{j=1}^{X(n)-\Phi_n} U_{(j)}, \tag{18}$$

where $U_{(1)} \le U_{(2)} \le \cdots \le U_{(M)}$ denote the order statistics of $U_1, \ldots, U_M$ as in Section 3.3. For n large enough we have $M \ge m$, and on the event $B_n \cap \{\Phi_n \le m\}$ we have

$$X(n+1) \ge \sum_{j=1}^{M-m} U_{(j)}. \tag{19}$$

Recall that $E(U_j) = e^{\alpha}$. From the first part of Lemma 3.4, we have that

$$E\Big[\sum_{j=M-m+1}^{M} U_{(j)}\Big] = O\Big(\sqrt{mM}\Big) = O\left(M\frac{n}{e^{(\alpha-\beta)n/2}}.\right)$$

Observe that $a_{n+1}e^{\alpha(n+1)} = \big(1 - \frac{2}{(n+1)^2}\big)M \cdot e^{\alpha}$ and that for large enough n we have that

$$\begin{aligned}
P\Big(\sum_{j=1}^{M-m} U_{(j)} < \big(1 - \tfrac{2}{(n+1)^2}\big)M \cdot e^{\alpha}\Big) &= P\Big(\sum_{j=1}^{M-m} U_{(j)} - E\Big[\sum_{j=1}^{M} U_j\Big] < -\tfrac{2Me^{\alpha}}{(n+1)^2})\Big) \\
&= P\Big(\sum_{j=1}^{M-m} U_{(j)} - E\Big[\sum_{j=1}^{M-m} U_{(j)}\Big] < -\tfrac{2Me^{\alpha}}{(n+1)^2} + E\Big[\sum_{j=M-m+1}^{M} U_{(j)}\Big]\Big) \\
&\le P\Big(\sum_{j=1}^{M-m} U_{(j)} - E\Big[\sum_{j=1}^{M-m} U_{(j)}\Big] < -\tfrac{Me^{\alpha}}{(n+1)^2}\Big).
\end{aligned} \tag{20}$$

By Chebyshev's bound (4) and the first part of Lemma 3.4,

$$\begin{aligned}
P\Big(\sum_{j=1}^{M-m} U_{(j)} - E\Big[\sum_{j=1}^{M-m} U_{(j)}\Big] < -\tfrac{Me^{\alpha}}{(n+1)^2}\Big) &\le \frac{(n+1)^4 \mathrm{Var}\Big(\sum_{j=1}^{M-m} U_{(j)}\Big)}{e^{2\alpha}M^2} \\
&\le \frac{(n+1)^4 \mathrm{Var}(U_1)}{e^{2\alpha}M} = O(e^{-\alpha n}).
\end{aligned}$$

Taking into account also (17) it follows that

$$\begin{aligned}
&P\big(X(n+1) \ge a_{n+1}e^{\alpha(n+1)} \mid B_n\big) \\
&\quad \ge P\big(X(n+1) \ge a_{n+1}e^{\alpha(n+1)} \mid \Phi_n \le m, B_n\big) P\big(\Phi_n \le m \mid B_n\big) \\
&\quad \ge P\Big(\sum_{j=1}^{M-m} U_{(j)} \ge a_{n+1}e^{\alpha(n+1)}\Big)\Big(1 - \frac{1}{(n+1)^2}\Big) \\
&\quad \ge \big(1 - O(e^{-\alpha n})\big)\Big(1 - \frac{1}{(n+1)^2}\Big) \ge 1 - \frac{2}{(n+1)^2},
\end{aligned}$$

for n large enough.

We proceed with the second and third probabilities on the right hand side of (16). We have, with V_j independent and having the distribution of V in Lemma 3.1, using that $E(V) = e^{\beta}$, (5) and that $Y(t) = \hat{Y}(t)$ whenever $X(t) > 0$,

$$\begin{aligned}
&P\big(\hat{Y}(n+1) < b_{n+1}e^{\beta(n+1)} \mid B_n\big) \\
&\quad = P\big(\hat{Y}(n+1) < b_{n+1}e^{\beta(n+1)} \mid X(n) \geq a_n e^{\alpha n}, b_n e^{\beta n} \leq \hat{Y}(n) \leq c_n e^{\beta n}\big) \\
&\quad = P\big(\hat{Y}(n+1) < b_{n+1}e^{\beta(n+1)} \mid b_n e^{\beta n} \leq \hat{Y}(n) \leq c_n e^{\beta n}\big) \\
&\quad \leq P\big(\hat{Y}(n+1) < b_{n+1}e^{\beta(n+1)} \mid \hat{Y}(n) = b_n e^{\beta n}\big) \\
&\quad = P\Big(\sum_{j=1}^{b_n e^{\beta n}} V_j < b_{n+1}e^{\beta(n+1)}\Big) \\
&\quad = P\Big(\sum_{j=1}^{b_n e^{\beta n}} V_j < \Big(1 - \frac{1}{(1+n)^2}\Big) b_n e^{\beta n} E(V)\Big) \\
&\quad \leq \frac{1}{b_n e^{\beta n}} \frac{(1+n)^4 \mathrm{Var}(V)}{E(V)^2} = O(e^{-\beta n}).
\end{aligned}$$

Similarly, but using (4) in place of (5) ,

$$\begin{aligned}
&P\big(\hat{Y}(n+1) > c_{n+1}e^{\beta(n+1)} \mid B_n\big) \\
&\quad \leq P\big(\hat{Y}(n+1) > c_{n+1}e^{\beta(n+1)} \mid \hat{Y}(n) = c_n e^{\beta n}\big) \\
&\quad = P\Big(\sum_{j=1}^{c_n e^{\beta n}} V_j > c_{n+1}e^{\beta(n+1)}\Big) \\
&\quad = P\Big(\sum_{j=1}^{c_n e^{\beta n}} V_j > \Big(1 + \frac{1}{(1+n)^2}\Big) c_n e^{\beta n} E(V)\Big) \\
&\quad \leq \frac{1}{c_n e^{\beta n}} \frac{(1+n)^4 \mathrm{Var}(V)}{E(V)^2} = O(e^{-\beta n}).
\end{aligned}$$

We conclude that (15) holds for n large enough. □

5 Proof of Theorem 1.4

The proof of Theorem 1.4 will be in two parts.

The case $\alpha < 0$. It is well known (see [9, Theorem 11.1]) that the probability that a subcritical branching process survives until time $t > 0$ decays

exponentially fast in t. That is, there exists $c > 0$ such that for every $t > 0$,

$$P(X(t) > 0) \le e^{-ct}.$$

Letting $T_X = \inf\{t : X(t) = 0\}$ we get that $E[T_{\rm u}] \le E[T_X] < \infty$. □

The case $0 < \alpha < \beta$. Similarly to (13) let $c = (\alpha + \beta)/2$ and $\delta = (\beta - \alpha)/4$, and note that $c - \delta = \alpha + \delta$ and $c + \delta = \beta - \delta$. We have that

$$\begin{aligned}
&P(T_{\rm u} \ge \tau) \qquad (21)\\
&\quad = P(T_{\rm u} \ge \tau, \hat{X}(t) \le e^{(c-\delta)t} < e^{(c+\delta)t} \le \hat{Y}(t),\ \forall t \ge \tau/2)\\
&\qquad + P(T_{\rm u} \ge \tau, \{\hat{X}(t) \le e^{(c-\delta)t} < e^{(c+\delta)t} \le \hat{Y}(t),\ \forall t \ge \tau/2\}^c)\\
&\quad \le P(T_{\rm u} \ge \tau, \hat{X}(t) \le e^{(c-\delta)t} < e^{(c+\delta)t} \le \hat{Y}(t),\ \forall t \ge \tau/2)\\
&\qquad + P(T_{\rm u} \ge \tau, \exists t \ge \tau/2 : \hat{X}(t) \ge e^{(\alpha+\delta)t} \text{ or } \exists t \ge \tau/2 : 0 < \hat{Y}(t) < e^{(\beta-\delta)t})\\
&\quad \le P(T_{\rm u} \ge \tau, \hat{X}(t) \le e^{(c-\delta)t} < e^{(c+\delta)t} \le \hat{Y}(t),\ \forall t \ge \tau/2)\\
&\qquad + P(\exists t \ge \tau/2 : \hat{X}(t) \ge e^{(\alpha+\delta)t} \text{ or } \exists t \ge \tau/2 : 0 < \hat{Y}(t) < e^{(\beta-\delta)t}).
\end{aligned}$$

For the first part of the right hand side of (21), we consider (for simplicity) first the case $\tau = 2n$, where we get

$$\begin{aligned}
&P(T_{\rm u} \ge \tau, \hat{X}(t) \le e^{(c-\delta)t} < e^{(c+\delta)t} \le \hat{Y}(t),\ \forall t \ge \tau/2)\\
&\quad = P(T_{\rm u} \ge 2n, \hat{X}(t) \le e^{(c-\delta)t} < e^{(c+\delta)t} \le \hat{Y}(t),\ \forall t \ge n)\\
&\quad \le P(0 < X(t) \le e^{(c-\delta)t} < e^{(c+\delta)t} \le Y(t),\ \forall t \in [n, n+1]) = P(A_n),
\end{aligned}$$

where A_n is as in Lemma 3.2. According to that lemma, there exists a $c_2 > 0$ such that for any n, we have that $P(A_n) \le e^{-2c_2 n} = e^{-c_2 \tau}$. It is easy to see that the same holds for all τ (adjusting c_2 if necessary).

For the second part of the right hand side of (21), we use Lemma 3.3, to conclude that there exists a $c_1 = c_1(\delta) > 0$ such that for any τ,

$$\begin{aligned}
&P(\exists t \ge \tau/2 : \hat{X}(t) \ge e^{(\alpha+\delta)t} \cup \exists t \ge \tau/2 : 0 < \hat{Y}(t) < e^{(\beta-\delta)t})\\
&\quad \le P(\exists t \ge \tau/2 : \hat{X}(t) \ge e^{\alpha t + \delta\tau/2}) + P(\exists t \ge \tau/2 : 0 < \hat{Y}(t) < e^{(\beta-\delta)t})\\
&\quad \le e^{-\delta\tau/2} + e^{-c_1\tau}.
\end{aligned}$$

We conclude that there exists $c_3 > 0$ such that $P(T_{\rm u} \ge t) \le e^{-c_3 t}$ for any $t > 0$, and so $E[T_{\rm u}] < \infty$. □

Remark 5.1. Clearly $E[T_{\rm u}] = \infty$ when $\alpha > \beta > 0$, since then $T_{\rm u}$ takes value ∞ with positive probability. We have not been able to determine in general whether or not $E[T_{\rm u}]$ is finite in the remaining case $\alpha = \beta$, but in the following special case it is easily seen to be finite. Suppose $\alpha = \beta = 0$, $\gamma_0 = 1$

and $\lambda = 0$. Then X and Y form independent critical branching processes. The extinction times T_X and T_Y for these respective processes satisfy

$$P(T_X > t) \sim \frac{1}{t}, \quad P(T_Y > t) \sim \frac{1}{t};$$

see [2, p. 159]. Thus $T_{\mathrm{u}} = \min\{T_X, T_Y\}$ satisfies

$$P(T_{\mathrm{u}} > t) = P(T_X > t)P(T_Y > t) \sim \frac{1}{t^2}$$

so

$$E[T_{\mathrm{u}}] = \int_0^\infty P(T_{\mathrm{u}} > t)dt \leq 1 + \int_1^\infty P(T_{\mathrm{u}} > t)dt \sim 1 + \int_1^\infty \frac{dt}{t^2} < \infty.$$

6 Applications of the main results

In this section we will briefly discuss some applications of our main theorems. Using our results on coexistence we are able to comment more on the issue of extinction of Y, which was the main focus of [3].

Central to the analysis in the present article were the auxiliary processes $\hat{X}$ and $\hat{Y}$. Recall that $\hat{Y}$ was in essence the process Y in an 'infinite sea of food', i.e. $X(0) = \infty$. However, if instead $X(t) = 0$, then $(Y(t+s))_{s\geq 0}$ has no healthy cells to feed on, and therefore $(Y(t+s))_{s\geq 0}$ grows at the exponential rate (see also (3))

$$\beta' = \bar{p} - 1 - \lambda = \alpha - \lambda.$$

The qualitative behavior of $(X(t), Y(t))_{t\geq 0}$ depends on the values of α, β and β'. We discuss the possible different regimes.

Regime 1. If $\alpha \leq 0$ then $(X(t))_{t\geq 0}$ eventually dies out, and since $\beta' \leq \alpha$, so does $(Y(t))_{t\geq 0}$. Hence $\eta = 1$.

Regime 2. If $0 < \alpha \leq \beta$ then if $\gamma_0 > 0$ it might be the case that $(Y(t))_{t\geq 0}$ dies out spontaneoulsy. However, if it does not, then according to Theorem 1.2, instead $(X(t))_{t\geq 0}$ will go extinct. If $\beta' \leq 0$, we then conclude that also $(Y(t))_{t\geq 0}$ dies out, that is $\eta = 1$. However, if $\beta' > 0$ then $(Y(t))_{t\geq 0}$ can survive on its own, that is $\eta < 1$.

Regime 3. If $0 < \beta < \alpha$ we are in the coexistence regime, in particular $\eta < 1$. As stated in Theorem 1.2, it might be the case that $X(t)Y(t) > 0$ for all $t > 0$. However, as in Regime 2, if $\gamma_0 > 0$, it is possible that $(Y(t))_{t\geq 0}$ dies out. Furthermore, if $(X(t))_{t\geq 0}$ dies out, then the behavior of $(Y(t))_{t\geq 0}$

would again be governed by the sign of β'.

Regime 4. If $\beta < 0$ then $(Y(t))_{t\geq 0}$ eventually dies out, that is $\eta = 1$.

We can draw qualitative conclusions from the above description, using also (3). For instance, if we fix $\alpha > 0$ and $E(\Gamma) \geq 1$ it follows that $\alpha \leq \beta$ for every $\lambda \geq 0$, and so we are always in Regime 2. As long as λ is small enough, so that $\beta' > 0$, the process $(Y(t))_{t\geq 0}$ can survive. This supports the intuition that small λ is good for the long term survival of $(Y(t))_{t\geq 0}$, see [3].

If instead $\alpha > 0$ while $E(\Gamma) < 1$ we see that we are in Regime 2 for small values of λ and in Regime 3 for large values of λ. Depending on the exact values of α and $E(\Gamma)$ we have the following possibilities:

- for small λ we have $0 < \alpha < \beta$, and $\beta' > 0$ so that $(Y(t))_{t\geq 0}$ might survive, that is $\eta < 1$;
- for slightly larger λ we can have $0 < \alpha < \beta$, and $\beta' \leq 0$ so that $(Y(t))_{t\geq 0}$ dies out, that is $\eta = 1$;
- for larger λ we have $0 < \beta < \alpha$, so that $(Y(t))_{t\geq 0}$ might again survive, that is $\eta < 1$;
- for even larger λ we have $\beta \leq 0$ so that $(Y(t))_{t\geq 0}$, again dies out, that is $\eta = 1$.

In [3], monotonicity of η as a function of λ was established when $\gamma_0 = 0$. In contrast, we see here that monotonicity of η in λ may fail if $E(\Gamma) < 1$ (and it is easy to find specific parameters for this to be the case). Note also the difference between the first case, in which $(Y(t))_{t\geq 0}$ is strong enough to survive on its own, and case three where $(Y(t))_{t\geq 0}$ *needs* the process $(X(t))_{t\geq 0}$ to feed on.

References

[1] B. C. Arnold and R. A. Groeneveld R.A. Bounds on expectations of linear systematic statistics based on dependent samples. *Mathematics of Operations Research*, 4(4):441–447, 1979.

[2] K. B. Athreya and P. Ney. *Branching processes*. Springer, 1972.

[3] J. E. Björnberg, E. I. Broman, T. Britton, and E. Nathan. A stochastic model for virus growth in a cell population. 2012.

[4] O. Garet and R. Marchand. Coexistence in two-type first-passage percolation models. *Ann. Appl. Probab.*, 15:298–330, 2005.

[5] E. J. Gumbel. The maximum of the mean largest value and the range. *Annals of Mathematical Statistics*, 25:76–84, 1954.

[6] O. Häggström and R. Pemantle. First passage percolation and a model for competing spatial growth. *J. Appl. Probab.*, 35:683–692, 1998.

[7] O. Häggström and R. Pemantle. Absence of mutual unbounded growth for almost all parameter values in the two-type richardson model. *Stoch. Proc. Appl.*, 90:207–222, 2000.

[8] T. E. Harris. A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc*, 56:13–20, 1960.

[9] T. E. Harris. *The theory of branching processes*. Springer, 1963.

[10] H. O. Hartle and H. A. David. Universal bounds for mean range and extreme observations. *Annals of Mathematical Statistics*, 25:85–89, 1954.

[11] C. Hoffman. Coexistence for richardson type competing spatial growth models. *Ann. Appl. Probab.*, 15:739–747, 2005.

[12] A. V. Kalinkin. Markov branching processes with interaction. *Russian Math. Surveys*, 57:241–304, 2002.